# A Bayesian Source Term inversion Method Based on Spatiotemporal Trajectory Prior and Joint Adaptive MCMC Sampling

Qing-Yun Li,[1,2,3] Jun-Fang Zhang,[1] Bing Lian,[1] Li-Ye Liu,[1] Rui Qiu,[2,3,*] and Jun-Li Li[2,3]

[1]*China Institute of Radiation Protection, Taiyuan 030006, China*
[2]*Department of Engineering Physics, Tsinghua University, Beijing 100084, China*
[3]*Key Laboratory of Particle & Radiation Imaging, Tsinghua University, Ministry of Education, Beijing 100084, China*

Determining the release source position and quantity is crucial for evaluating the consequences of atmospheric radionuclide release events, with the Bayesian method serving as the primary tool for source inversion. Reducing the impact of input data errors on inversion uncertainty and improving computational efficiency are key to developing robust and efficient inversion algorithms. To address these challenges, we developed a spatiotemporal trajectory prior (STP) distribution that effectively mitigates the influence of measurement and simulation errors on inversion results without increasing computational costs, thereby enhancing the robustness and accuracy of the inversion process. Additionally, we introduced a joint adaptive Markov Chain Monte Carlo (MCMC) sampling method that integrates the traditional parallel tempering (PT) algorithm with a novel joint adaptive transition proposal (JATP) algorithm to accelerate inversion calculations. The proposed methods were optimized and validated using data from the first release of the European Tracer Experiment (ETEX-I). After determining the hyperparameters, the JATP algorithm consistently maintained the sampling process near the theoretically optimal acceptance rate of 0.234. The PT algorithm, utilizing an optimized temperature schedule, achieved a 2.89-fold improvement in sampling efficiency compared to single-chain sampling. Under bootstrap statistical comparison, the method reduced the relative error of position, relative error of release quantity, and total relative error by 25.9%, 27.7%, and 27.8%, compared to the traditional uniform prior method, respectively. And the deviation of the estimated and true source position is within 0.25. The results demonstrate the accuracy and effectiveness of our method.

Keywords: Radionuclide source inversion. Spatiotemporal trajectory prior distribution. Joint adaptive transition proposal.

## I. INTRODUCTION

Environmental protection is a critical consideration in the development of the nuclear industry. Various stages, including the operation of nuclear power plants, isotope production, and the reprocessing of radioactive waste, involve environmental concerns related to the atmospheric release of radionuclides. Since the Chernobyl nuclear accident[1] and the Fukushima accident[2], this issue has garnered significant attention from both the academic community and the general public[3]. In recent years, several incidents involving radionuclide releases have further highlighted these concerns, including the abnormal detection of $^{131}$I in 2011[4], the $^{106}$Ru leak in 2017[5], and the detection of $^{134/137}$Cs in 2020[6]. These events underscore the challenges of accurately predicting the dispersion of radioactive pollutants, primarily due to the lack of crucial source term information, such as the position, timing, and quantity of releases. This knowledge gap complicates the evaluation of accident consequences and hampers effective emergency response. Consequently, the accurate and timely inversion of source term information has become an essential priority[7, 8].

Methods for source inversion are generally divided into two categories: optimization methods and probabilistic modeling methods. Research on optimization methods[9–12] primarily focuses on enhancing cost function constraints through regularization techniques[13–15] and developing more efficient optimization algorithms[14, 16]. However, these methods often struggle with accurately identifying the source position, prompting the integration of auxiliary positioning algorithms in some studies[17–20]. Probabilistic modeling methods, typically framed within Bayesian inference, construct the posterior distribution of source term parameters based on model-observation hypotheses[21, 22]. This approach provides statistically robust and interpretable results by delivering distributions of reconstructed parameters. Recent studies demonstrate the effectiveness of Bayesian methods in source inversion[23–29]. Nonetheless, most previous work has focused on designing likelihood functions using the L1 or L2 norm of the difference between simulated and observed detections, often overlooking the critical role of the prior function[23–26, 29]. This neglect limits the effective use of sparse detection data, exacerbating the inherent ill-posedness of inversion problems. Real-world scenarios further complicate this process, as detection data are often sparse and of varying quality. Additionally, inaccuracies in meteorological input fields and errors in dispersion model simulations can unpredictably affect inversion results, posing significant challenges to achieving robust source term estimates using Bayesian methods[7, 22]. Moreover, deriving accurate posterior distributions for source parameters frequently requires extensive iterative sampling, making computational efficiency a critical concern for Bayesian probabilistic modeling[30–32].

To address the limitations of traditional Bayesian methods in complex scenarios—namely, insufficient robustness, accuracy in source inversion, and slow computational speed—we propose an enhanced Bayesian framework that integrates the spatiotemporal trajectory prior (STP) distribution and a joint adaptive Markov Chain Monte Carlo (MCMC) sampling ap-

* Corresponding author, qiurui@tsinghua.edu.cn.

proach for source inversion. The STP distribution is founded on the principle that, for a single release source, all plumes passing through detection points originate and overlap at the source location. This provides a robust prior for source term parameters, incorporating additional information beyond what L1 or L2 norm distributions can offer. Consequently, it significantly improves the accuracy and robustness of source inversion results. Moreover, the STP approach allows for a rapid and reasonably accurate estimation of the potential range of source term parameters prior to complex iterative computations, making it broadly applicable to various inversion algorithms beyond the Bayesian framework. To further enhance computational efficiency and ensure stable convergence of MCMC sampling, we proposed a joint adaptive sampling method based on the classical parallel tempering MCMC approach. This method combines the adaptability of single adaptive transition proposals for high-dimensional complex distributions with mechanisms to mitigate deviations from Markovian properties during the sampling process. As a result, it ensures efficient and reliable sampling of high-dimensional distributions, addressing key challenges in source inversion under complex conditions.

Using data from the first release of the European Tracer Experiment (ETEX- I)[33], we designed three validation schemes to optimize the hyperparameters of the proposed method and validate its effectiveness, as well as its sensitivity to input errors. In previous studies, researchers typically used the data from fixed-sites to validate inversion methods[7, 29]. However, Xu et al..[7] highlight that the choice of sites has a significant impact on inversion results. Therefore, we employed the bootstrap method[34] to statistically evaluate the properties of the proposed method, providing a robust validation framework.

## II. MATERIALS AND METHODS

This section first presents the modeling approach for the atmospheric radionuclide release source inversion problem. It then introduces the construction of the STP function and likelihood function within the Bayesian framework, followed by a discussion on the joint adaptive MCMC sampling method. Finally, the section describes the validation benchmarks and evaluation metrics for the proposed method.

### A. Modeling

When solving inversion problems, we typically know the detection vector $\boldsymbol{Y}$, and attempt to solve for the release vector $\boldsymbol{X}$. Inverse modeling usually assumes that these vectors satisfy a mapping relationship,

$$\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{X} + \varepsilon \tag{1}$$

Where $\boldsymbol{H}$ is denoted as the source-receptor sensitivity matrix, hereafter referred to as SRSM, and $\varepsilon$ represents the possible error vector, which is explained in more detail in relevant study[35].

In this paper, the release time of release source is considered, as it is usually unknown in most cases that require estimation. In inversion efforts involving unknown position release source, it is common practice to assume that the source is a homogeneous release at a specific point in time to reduce the dimensionality of the solution space[7, 16, 36],i.e., the source term is fully characterized by parameter $\boldsymbol{X} = [\boldsymbol{x}, t_s, t_e, Q]$, where $\boldsymbol{x} = (x_{lon}, x_{lat})$ represents coordinates of source release position, and $t_s$, $t_e$ and $Q$ represent the start time, end time, and total release of the release source, respectively. Thus, eq.1 can be expressed as:

$$\begin{aligned} y_j &= \widehat{y}_j + \varepsilon_j \\ &= \sum_k h_{jk}\left(\boldsymbol{x_i}\right) \frac{Q \Delta t_k}{t_e - t_s} + \varepsilon_j \end{aligned} \tag{2}$$

where, $h$ represents the elements in the transfer matrix $\boldsymbol{H}$, $j = 1 \cdots m$ with $m$ is the number of receptor samples, $i = 1 \cdots n$ with $n$ represents the number of possible release positions and $k = 1 \cdots l$, where $k$ denotes the number of discrete intervals ($\Delta t_k$) between $t_s$ and $t_e$. The error between detected data $y_j$ and simulated data $\widehat{y}_j$ is denoted as $\varepsilon_j$, which includes both measurement error and simulation error. It is often difficult to assess this error in practice, so we estimate the covariance matrix $\boldsymbol{R}$ of the detected and simulated datas as an unknown parameter in the source term inversion, i.e., $\boldsymbol{X} = [\boldsymbol{x}, t_s, t_e, Q, \boldsymbol{R}]$.

The source-receptor sensitivity matrix (SRSM) describes the sensitivity of observations with respect to the unit release. Rajaona et al. [37] noted that this is the most time-consuming step in the inversion algorithm. Fortunately, the adjoint-based backward dispersion computation addresses this issue and has been demonstrated by Seibert et al. [38] to be equivalent to forward dispersion. In source term inversion problems with unknown position, the number of receptor points is significantly smaller than the potential source positions, making it well-suited for calculating the SRSM matrix through backward dispersion[39]. In this work, we perform dispersion calculations using the backward mode of the Lagrangian Particle Dispersion Model (LPDM) FLEXPART [40].

### B. Probabilistic description

The implementation of the Bayesian inversion algorithm relies on the application of Bayes' rule:

$$p\left(x|y\right) = \frac{p(y|x)p\left(x\right)}{p\left(y\right)} \propto p(y|x)p\left(x\right) \tag{3}$$

with the objective of obtaining the posterior distribution $p(x|y)$ for the scenario by estimating a reasonable likelihood function $p(y|x)$ and prior distribution $p(x)$. Constructing the posterior probability density function is crucial for the accuracy of source term inversion. A well-designed construction ensures that the inferred distribution of source parameters in complex scenarios aligns more closely with the true source term.

### 1. Prior distribution of variables

It is common practice to assume that the variables are independent of one other in sampling or fitting algorithms for multivariate coupled complex probability density function[23–29]. Under this assumption we set priors for each variable separately and define the product as the total prior:

$$P(\boldsymbol{X}) = P(\boldsymbol{x_i}) \cdot P(t_s) \cdot P(t_e) \cdot P(Q) \cdot P(\boldsymbol{R}) \quad (4)$$

In almost all previous work, the prior probability density functions for release location and time were chosen as uniform distributions over the range of possibilities due to the lack of prior knowledge[23–26, 29, 41–43]. Inspired by previous studies[15, 18], we have developed a backward trajectory-based computation of spatiotemporal prior distributions, which provides prior information that has not been utilized in previous probabilistic modeling algorithms, thereby enabling a more accurate estimation of the source term.

As mentioned in section II A, SRSM is computed in this work using the backward dispersion model. This calculation incorporates the backward dispersion trajectory information with the receptor points serving as the source, which is required for the STP. As a result, the inclusion of the STP does not necessitate additional dispersion computations. The backward dispersion plume results represent the sensitivity of the plume's coverage of spatiotemporal coordinates to the receptor. In other words, if a spatiotemporal coordinate represents the real release source, the concentration detected at the corresponding receptor site can be expressed as:

$$conc_{rec} = Q_{rel}\frac{s}{\Delta T v} \quad (5)$$

where $conc_{rec}$ is the concentration at the receptor site, $Q_{rel}$ is the release from source, $s$ is sensitivity to emission from the source, and $\Delta T$, $v$ represent the time interval and grid volume of simulation, respectively. Clearly, when a concentration of radioactive contaminants is obtained at a receptor site ($conc_{rec} \neq 0$), the source must be within the coverage of the backward plume ($s \neq 0$). Conversely, when concentration is not detected at a receptor site ($conc_{rec} = 0$), the releasing source ($Q_{rel} \neq 0$) must be outside the coverage of backward plume ($s = 0$).

Based on the above logic, we extract prior information from backward plumes. For each backward plume matrix (hereafter referred to as $\boldsymbol{BPM}$, it is a three-dimensional matrix that stores the results of backward dispersion sensitivity), the sensitivity to emission decreases sharply with increasing diffusion distance, usually differing by several orders of magnitude. Therefore, in this study, we characterize the backward plume at receptor sites with detection values using weights $K$, and define the following spatiotemporal trajectory matrix:

$$\boldsymbol{M}_{>0}(\boldsymbol{x_i}, t) = \sum_{i=1}^{N_{>0}} K \; if \; \boldsymbol{BPM_i}(\boldsymbol{x_i}, t) \neq 0 \quad (6)$$

where $N_{>0}$ is number of receptors with non-zero measurements.

In the early stages of an accident, there are typically only a few detections at selected receptors, while the majority of receptors have no detections. However, this part of the information is still of great value. Considering the detection limit of the detection equipment in practice, it may not be possible to detect very low concentrations of radioactive contaminants. Therefore, the possibility that the spatiotemporal coordinates covered by backward plumes of non-detected receptors cannot be completely ruled out must be acknowledged. After considering non-detection receptors, we correct spatiotemporal trajectory matrix as:

$$\boldsymbol{M_{ST}} = \boldsymbol{\lambda} \circ \boldsymbol{M_{>0}} \quad (7)$$

$$\boldsymbol{\lambda}(x_i, t) = 1 - \frac{\sum_{j=1}^{N_0} 1 \; if \; \boldsymbol{BPM_j}(\boldsymbol{x_j}, t) \neq 0}{N_0} \quad (8)$$

where $N_0$ is the number of non-detection receptors, $\circ$ represents Hadmard product.

The source is necessarily located in the spatiotemporal region where the backward trajectories overlap most significantly. We define the temporal prior as:

$$P_t(t) = normalize(\; \max_{\boldsymbol{x}} \boldsymbol{M_{ST}}(\boldsymbol{x}, t) \;) \quad (9)$$

Here, $normalize$ refers to the process of adjusting the probability density so that its integral equals one.

Since the solution variables in our work are the start time and the end time, the corresponding prior distributions are treated as follows:

$$P(t_s) \cdot P(t_e) = \frac{\sum_{t \in [t_s, t_e]} p_t(t)}{\max(p_t(t)) \cdot (t_e - t_s)} \quad (10)$$

Similarly, we define the spatial prior, but considering that the extent of the backward plume gradually increases with diffusion time—leading to significant artifacts in the spatial prior—we weight it using the temporal prior function:

$$P_s(\boldsymbol{x}) = normalize(\; \max_t(\boldsymbol{M_{ST}}(\boldsymbol{x}, t) \cdot P_t(\mathrm{t})) \;) \quad (11)$$

For the choice of $K$,we have two approaches: 1) the adoption of uniform weights, and 2) weighting with the detection values of receptors. The information corresponding to method 2 is implicit in the intrinsic properties of likelihood function. Therefore, method 1 is chosen in this work to obtain additional trajectory overlap information, which further constrains the Bayesian inversion method. The spatiotemporal prior effectively identifies the high-probability density regions of the source term in both time and space, and it has been shown to align well with the true source term (see section III A 1 for details). Notably, the spatiotemporal prior can be generalized to nearly all inversion methods, as it quickly determines approximate information about the source term.

For total release quantity $Q$, given that the range of values may span multiple orders of magnitude, the logarithm of $Q$ is assumed to follow a uniform distribution. Saunier et al.[18]

demonstrated that the diagonal matrix assumption is justified for detections that are far apart, meaning the detections are independent of each other. For the sake of simplification in calculations, we assume that the accuracy of the detecting instruments is comparable, thus supporting the homoscedasticity assumption ($\boldsymbol{R} = \boldsymbol{I}r$) . In this work, we assume that $r$ follows a uniform distribution over a specified range (0.01-20).

### 2. *Likelihood*

Likelihood is a quantitative description of the probability of detections given a source term. Specifically, it is a function that quantifies the difference between the vector of detected values $\boldsymbol{Y_{\text{det}}}$ and the vector of simulated values $\widehat{\boldsymbol{Y}}$ of the source term to be determined.

Dumont Le Brazidec et al.[26] compared the performance of different likelihood functions and points out that the choice of likelihood function significantly impacts the performance of the inversion model. Referring to previous studies, the log-normal distribution may be the better choice for a likelihood function:

$$P(\widehat{\boldsymbol{Y}}|\boldsymbol{X}) = \frac{e^{-\frac{1}{2}\ln\left(\frac{\widehat{\boldsymbol{Y}}}{\boldsymbol{Y_{\text{det}}}}\right)^T \boldsymbol{R}^{-1}\ln\left(\frac{\widehat{\boldsymbol{Y}}}{\boldsymbol{Y_{\text{det}}}}\right)}}{\sqrt{(2\pi)^{N_{\text{det}}}}|\boldsymbol{R}|^{1/2}\prod_{i=1}^{N_{\text{det}}} y_i} \tag{12}$$

where $N_{\text{det}} = N_{>0} + N_0$ and $\widehat{\boldsymbol{Y}}$ is calculated from eq.1 i.e. $\widehat{\boldsymbol{Y}} = \boldsymbol{HX}$. This study utilizes many zero-value detections (detections under detection limit), which fall outside the domain of definition for the logarithm. Therefore, we set a truncation threshold, as in the work of Liu et al.[44] and define a function in place of the logarithmic function in eq12:

$$\zeta(y) = \ln\left(\frac{y+\theta}{y_{ref}}\right) \tag{13}$$

where $y_{ref}$ is defined as reference concentration and $y_{ref} = 2\theta = \min(y_i)$ which references the work of Dumont Le Brazidec et al.[25].

## C. Joint adaptive MCMC sampling

MCMC is a class of algorithms used to sample from the probability distribution $P(\boldsymbol{X}|\boldsymbol{Y})$. The core idea is to construct a Markov chain with the target distribution as its equilibrium distribution. By simulating the Markov chain and running it for a sufficient number of steps, the state of the chain will eventually converge to the target distribution. The Metropolis-Hastings (MH) algorithm is one of the most widely used methods in MCMC, with its sampling efficiency largely dependent on the choice of the transition proposal function[23]. The adaptive transition proposal function addresses the issue of step size adaptation for arbitrarily complex target distribution, but it also introduces non-Markovian behavior into the sampling process. In this study, we proposed a joint adaptive transition proposal (JATP) method to mitigate this problem, ensuring robustness of the sampling process. Specifically, we use three adaptive Metropolis proposal functions: Adaptive Metropolis (AM)[45], Differential Evolution-Markov Chain (DE-MC) algorithm[46], and Single Component Adaptive Metropolis (SCAM)[47], proportionally in the sampling process. In practice (see section III A 2 for details), this method ensures that the converged acceptance rate approaches the theoretically derived optimal rate (approximately 0.234) for 1-10 dimensional cases, as suggested by Roberts et al.[48].

When sampling high-dimensional probability density functions, MCMC is often combined with parallel tempering (PT) algorithms to balance local characterization and global exploration. [49]. The choice of temperature schedule is critical. Many of the previous studies suggested that a geometric spacing was optimal[49], therefore, in this work, we similarly choose geometric spacing and investigate the optimal temperature schedule (see section III A 3 for details) selection in the source term inversion problem. We finalized the temperature schedule using 8 Markov chains with different temperatures, following a geometric progression. The rate $\lambda$ is set to the natural logarithm base $e$ , such that $T_{i+1} = \lambda T_i$, where $i = 1, 2, \ldots, 7$, with $T_1 = 1$ and $T_8 = e^7 \approx 1097$. Inter-chain exchanges are performed every 100 steps.

## D. Evaluation methodology

### 1. *Evaluation metrics*

1. *source inversion.* Relative errors in position ($\delta_x$) total release ($\delta_Q$) and total relative errors ($\delta_{total}$) are defined to assess the accuracy of the reconstructed source terms and the reliability of the method:

$$\delta_{\boldsymbol{x}} = \frac{|\boldsymbol{x_{true}} - \boldsymbol{x_{rec}}|}{L} \times 100\% \tag{14}$$

$$\delta_Q = \frac{|Q_{true} - Q_{rec}|}{Q_{true}} \times 100\% \tag{15}$$

$$\delta_{total} = \sqrt{\delta_x^2 + \delta_Q^2} \times 100\% \tag{16}$$

where $\boldsymbol{x_{true}}$ and $Q_{true}$ represent the position and total release of the real source term, respectively, while $\boldsymbol{x_{rec}}$ and $Q_{rec}$ represent the position and total release of the reconstructed source term. $L$ denotes the characteristic length of the simulation domain (the longer boundary of the experimental domain is used). The total relative error is the $L_2$ paradigm of each error component. Since it is difficult to accurately measure the inversion accuracy of the source item nuclide release time intervals by the statistical values of the release start and end times, only individual inversion cases are illustrated without setting statistical metrics.

2. *Jensen-Shannon (JS) divergence*. JS divergence is a common measure of the similarity between two probability density distributions[50] and is defined as follows:

$$D_{JS}P||Q = \frac{1}{2}D_{KL}(P||M) + \frac{1}{2}D_{KL}(Q||M) \tag{17}$$

$$D_{KL}(P||Q) = \sum\nolimits_{i} P(i) \log \frac{P(i)}{Q(i)} \tag{18}$$

where $M$ is the mean distribution of the distributions $P$ and $Q$, i.e., $M = \frac{1}{2}(P+Q)$. The value of JS divergence is in the range of $[0, 1]$, the closer the JS divergence is to 0 the more similar the two distributions are, and conversely, the further from 0, the more different they are.

### 2. *Statistical method*

The bootstrap method was employed for statistical validation of the effectiveness of the proposed approach. It is a computationally convenient statistic (parameter estimation or statistical inference for sample sets with multiple repetitive sampling) that does not depend on assumptions about the form of the distribution or the variance, and remains valid for small samples. It is much more accurate than classical inferences based on Normal or t distributions in many cases.

### 3. *Validation experiment*

The proposed method utilizes the ETEX-I field experiment as a benchmark for optimization and validation.

The ETEX-I experiment, conducted from October to November 1994, was an atmospheric tracer study organized by several European countries, involving a total of thirty-six organizations worldwide[33]. The experiment involved a uniform tracer release at Monterfil in Brittany, France (48.058° N, 2.0083° W), occurring from 16:00 UTC on October 23 to 03:50 UTC on October 24. During this release period, a total of 340 kg of PMCH was emitted into the atmosphere. The experiment yielded 3,104 detections from 168 ground sites, with each data representing a three-hour average concentration. The release position and the locations of the ground sites are shown in Fig. 1a.

### 4. *Settings of validation schemes*

1. *fixed-sites validation experiment*. Detecting data from a randomly selected set of ground stations (B03, F03, N01, R02) was utilized to optimize the joint adaptive MCMC sampling method, specifically for hyperparameter selection.The positions and labels of the selected stations are depicted in Fig. 1b. For the JAPT algorithm, we separately conducted 50 iterative sampling calculations using AM, DE-MC, and SCAM individually, comparing the acceptance rate curves of each scheme to determine the weight ratio of the basic algorithms in the joint adaptive approach. For the temperature schedule of the PT algorithm, we compared different temperature schemes based on their burn-in periods to identify the relatively optimal configuration. The joint adaptive MCMC sampling method, with these optimized hyperparameters, was then applied to subsequent source term inversion calculations.

2. *statistical validation experiment*. Using the bootstrap method, random sampling with replacement was conducted from all detecting data of ETEX-I to explore the properties of STP. Taking the STP generated from the full dataset as a reference, the sensitivity of STP to data

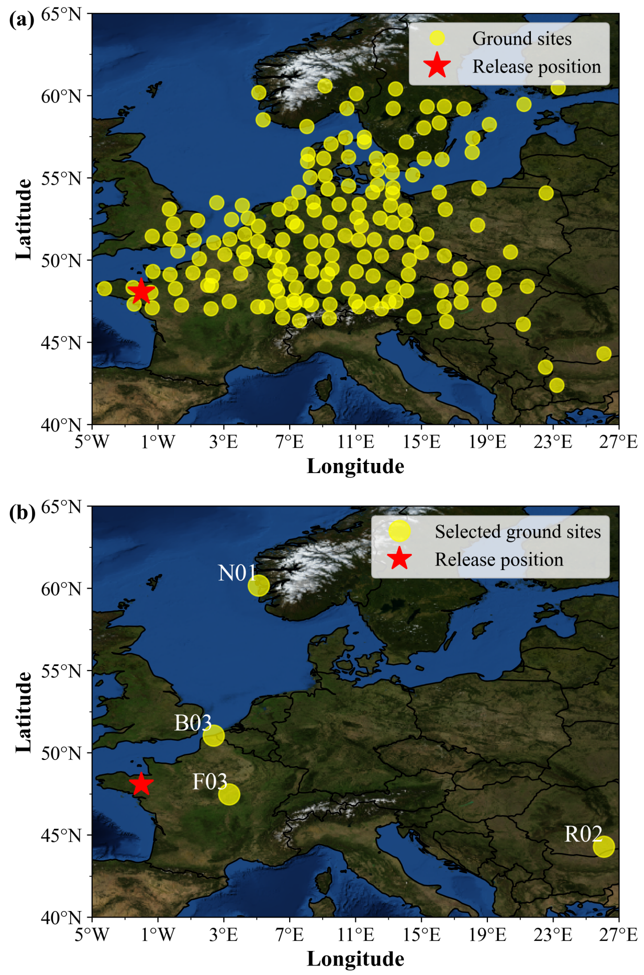


Fig. 1. Positions of release source and ground sites. The red star indicates the true source release position, while the yellow dots represent the positions of ground detection sites. (a) Positions of source and all ground sites in ETEX-I. (b) Positions of the randomly selected ground sites (B03, F03, N01, R02). The basemap is sourced from NASA's October Blue Marble Next Generation, featuring topography and bathymetry (https://visibleearth.nasa.gov/collection/1484/blue-marble).

volume is assessed through a comparison of JS divergence. Based on these results, the required data volume for inversion calculations is determined, enabling a statistical analysis of the improvements introduced by the STP algorithm to the inversion problem.

3. *ideal twin experiment*. In practical cases, it is challenging to estimate the errors in input data. Therefore, using the release scenario from ETEX-I, model simulation results are treated as true concentration values. By introducing relative errors, a twin experiment is designed to investigate the sensitivity of the Bayesian inversion method based on STP to input data errors under complex scenarios.

## III. RESULTS AND DISCUSSION

### A. Hyperparameter optimization for joint adaptive MCMC method

#### *1. The inversion performance of fixed-sites scheme*

Table 1 compares the inversion results of the fixed-sites scheme used for parameter optimization with the true source term, revealing that the inversion results are quite good, with a location error of 0.5% (15.71 km) and a relative error of 26% for the release quantity. While this result is generally satisfactory, the inversion of the release time indicates that, although the end time is accurately reconstructed, the start time is relatively earlier. This trend aligns with the temporal prior and is a consequence of the backward dispersion of the "plume" (described in Section 3.2.1).

#### *2. Weight selection for JATP algorithm*

Through comparative testing, we determined the optimal mixing ratio ($DE$:$AM$:$SCAM$ = 0.56:0.22:0.22) of the three adaptive transition proposal algorithms. As shown in Fig. 2, the acceptance rate of the DE algorithm gradually increases with the number of iterations, stabilizing around 0.3—higher than the optimal acceptance rate. In contrast, the acceptance rates for the AM and SCAM algorithms gradually decrease, converging to 0.06 and 0.19, respectively, ultimately stabilizing at rates lower than optimal. Notably, the joint adaptive transition proposal algorithm mitigates the non-Markovian nature introduced by the historical trajectories of single adaptive approaches, ensuring that the acceptance rate remains near the theoretical optimum throughout the process, even from the early stages of iteration. This indicates that the JATP algorithm achieves an optimal match between step size and the geometric characteristics of the high-dimensional target distribution, thereby attaining superior sampling efficiency[48].

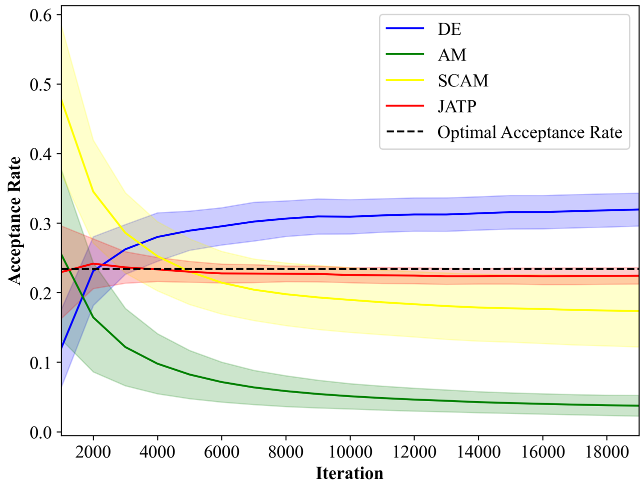


Fig. 2. Curves depicting the acceptance rates against the number of iterations for various adaptive transition proposal algorithms, based on 80 runs. The solid lines represent the average acceptance rates every 50 iterations for each algorithm, while the dashed line indicates the optimal acceptance rate (0.234). The shaded areas represent one standard deviation around the mean.

#### *3. Temperature schedule of PT algorithm*

Fig. A.1 illustrates the convergence curves of source term variables from multiple runs with varying numbers of tempering chains, all maintained at a fixed geometric spacing $e$. While all curves converge near the true value, the convergence iterations for the parallel tempering MCMC occur significantly earlier than those of the classical single-chain MCMC. Furthermore, as the number of chains increases, the iterations for variable convergence consistently occur earlier. Notably, unstable convergence was observed at $n = 5$ , which may be attributed to the intrinsic properties of the sampled distribution. Fig. 3(a) presents the average burn-in period for each temperature schedule under these conditions. It is evident that after $n = 3$ (i.e., two tempering chains with a maximum temperature of $e^2$), the rate of convergence stabilizes. Figs. A.2 and A.3 display the trace plots for multiple sampling calculations with a fixed number of chains ($n = 3$) and a fixed maximum temperature ($T_{\max} = e^2$), respectively. Figs. 3(b) and 3(c) show the corresponding average burn-in period variation curves. As illustrated in Fig. 3(b), when the number of chains is fixed, an increase in geometric spacing initially causes a decrease in the convergence position or iteration of the variable, followed by an increase. The optimal geometric spacing is observed at $e \sim 3.2$, as excessive geometric spacing ($> 3.2$) can hinder exchanges between adjacent chains, leading to insufficient exploration of the solution space by the low-temperature Markov chain and difficulty in escaping local extrema. Conversely, if the geometric spacing is too small ($< e$), although chain exchanges occur more frequently, the relatively low maximum temperature fails to fully leverage the global search capability of the high-temperature chains. This results in minimal gains from

Table 1. Source inversion result and relative errors for selected ground sites (B03, N01, F03, R02). $\boldsymbol{X_{rec}}$ and $\boldsymbol{X_{real}}$ stand for the inversion result and the true source term, respectively.

| | Parameters | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Longitude | Latitude | $Q_t$ | $t_s$ | $t_e$ | $\delta_x$ | $\delta_Q$ | $\delta_{total}$ |
| $\boldsymbol{X_{rec}}$ | $-2.0316°E$ | $47.919°N$ | 429.8kg | 23T4:30 | 24T4:00 | 0.005 | 0.26 | 0.26 |
| $\boldsymbol{X_{real}}$ | $-2.0083°E$ | $48.058°N$ | 340kg | 23T16:00 | 24T3:50 | | | |

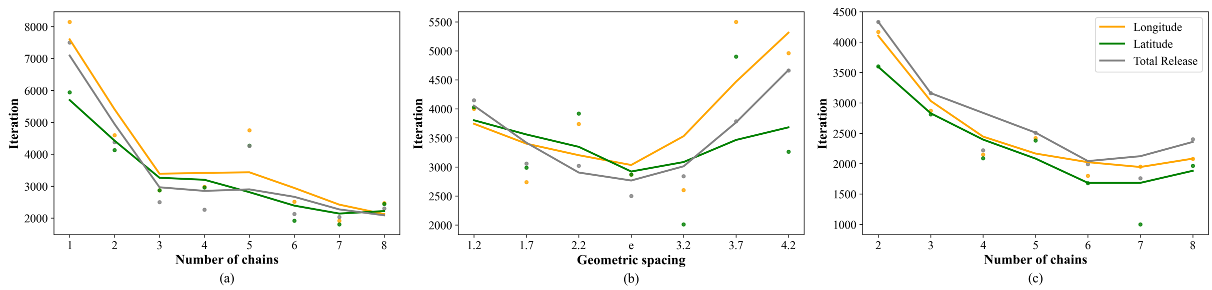


Fig. 3. LOWESS regression curves of the average burn-in period under various temperature schedules. (a) The geometric spacing is fixed at $e$; (b) The number of parallel chains is fixed at 3; (c) The maximum temperature is fixed at $e^2$.

each exchange while adding computational burden without significantly enhancing exploration performance. Fig. 3(c) demonstrates that, with a fixed maximum temperature, the number of iterations required for convergence decreases as the number of chains increases. However, once the number of chains reaches 4 or more, the convergence iterations tend to stabilize. This indicates that when the maximum temperature is fixed, as long as the geometric spacing between adjacent annealing chains meets the exchange requirements, the sampling performance of the Markov chain will remain consistent. Based on simulation comparisons, the geometric spacing of MCMC sampling in our inversion problem should be between $e$ and 3.2, with the maximum temperature exceeding $e^2$. Consequently, we selected a temperature schedule featuring 8 chains (7 tempering chains) with a geometric spacing of $e$ and a maximum temperature of $1097(e^7)$. This schedule resulted in an approximate 2.89-fold improvement in convergence speed compared to the single-chain MCMC method.

### B. Source inversion based on STP

#### 1. *Effectiveness of STP*

Fig. 4 presents the prior probability density distribution of the STP, derived from the all ground-based detection data collected during the ETEX-I experiment. To enhance clarity, we have excluded the negligible probability segments (those below a specified threshold $1e^{-6}$) in the spatial probability density plot Fig. 4a. Notably, the true source position—indicated by a star symbol in Fig. 4a—and the actual release time, represented by the interval between the vertical lines in Fig. 4b, align closely within the region of maximum STP, reinforcing the robustness of our method for characterizing the source term. However, we observe a tendency for the region of maximum prior probability density to shift to an earlier position (west of the source in Fig. 4a, aligned with the mean prevailing wind direction from west to east), and a similar pattern is evident in Fig. 4b. This shift arises from the sensitivity computations in the SRSM, which fundamentally follows a backward diffusion process. As the plume expands gradually during the diffusion, the decay of plume overlap becomes slow after reaching an extreme value. This gradual decay results in an artifact of elevated probability density in the STP estimation during earlier moments. Nevertheless, this effect does not compromise the validity of the method, as detailed in section III B 2. Fig. 5 illustrates the impact of using the temporal prior to weight the spatial prior, as mentioned in section II B 1. After applying the temporal prior to the spatial prior (Fig. 5b), it becomes clear that the distribution of probability densities is more concentrated around the actual release source compared to the unweighted distribution (Fig. 5a). As shown in Fig. 5c, the significance of the temporal weighting lies in its ability to increase the probability density near the true source while decreasing it in other regions, which, according to eq.3, effectively modifies the posterior probability density distribution.

#### 2. *Sensitivity of STP to data volume*

Fig. 6 illustrates the distribution of JS divergence between the STP computed with varying numbers (10, 20, 30, 40, 50, and 60) of detections and the STP derived from the full dataset. For each data volume, 300 ($3104/10 \approx 300$) independent computations were performed. The JS divergence of the spatial prior distribution is slightly higher than that of the temporal prior distribution. This discrepancy arises because different site selections lead to variations in the backward plume release positions, resulting in differences in the spatial prior distribution, particularly near the chosen sites. However, this effect diminishes closer to the source as the

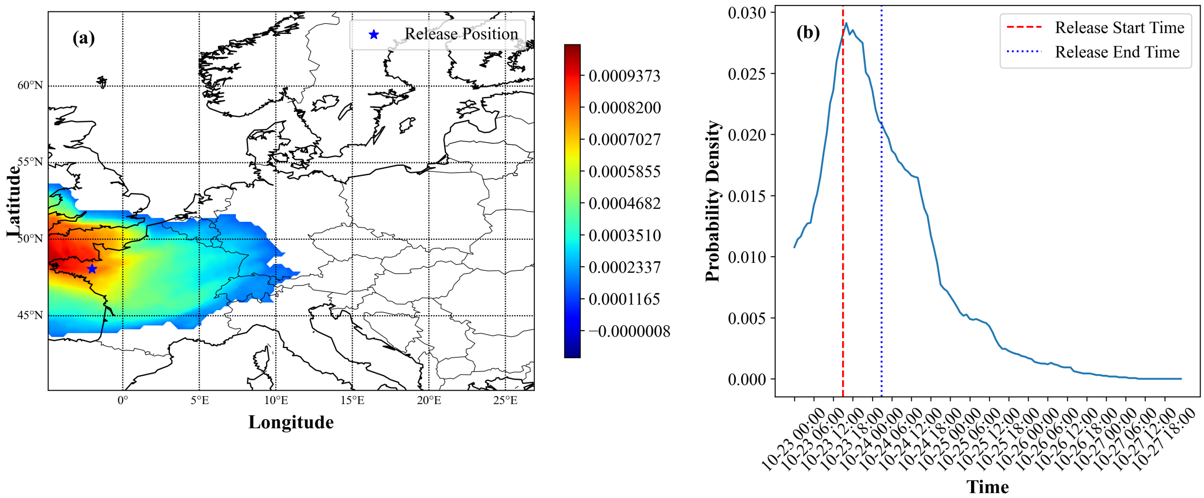


Fig. 4. STP probability density distribution derived from all detections. (a) Spatial probability density distribution; (b) Temporal probability density distribution.

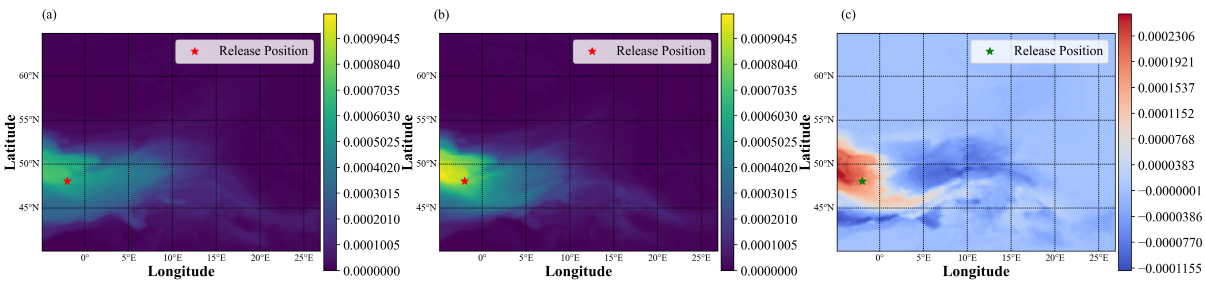


Fig. 5. Impact of temporal prior weighting on spatial prior. (a) Spatial prior unweighted by temporal prior. (b) Spatial prior weighted by temporal prior. (c) Comparison of changes before and after weighting.

plume develops over time. In contrast, the sensitivity to emissions represented by the backward plumes is maximized at the release time due to the uniqueness of the actual release source, leading to much less variation in the temporal prior. Consequently, the mean JS divergence lies between the two distributions. As the number of detections increases, the interquartile range (IQR) of JS divergence remains below 0.1 when the data volume exceeds 30. As the number of detections increases, the interquartile range (IQR) of JS divergence remains below 0.1 when the data volume exceeds 30, indicating that STP demonstrates robust performance in estimating the source term distribution even with a limited quantity of detection data. In this study, considering the scarcity of real-scenario detections and the need for the prior function to be as accurate as possible, we set the number of detections for each subsequent computational simulation at 40. Consequently, the number of independent computations used for statistical analysis was 80 ($3104/40 \approx 80$).

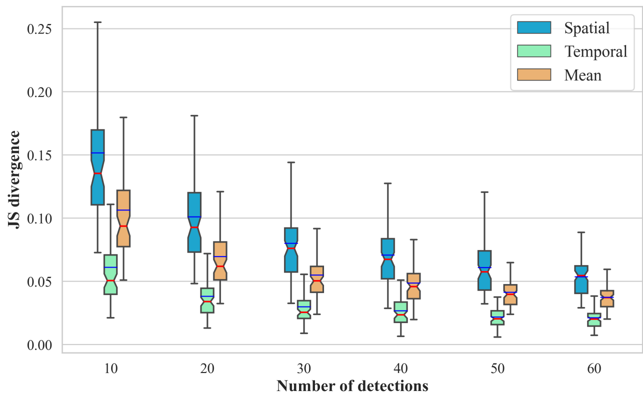


Fig. 6. Sensitivity of the spatiotemporal trajectory prior to site selection and data volume. The red solid line represents the median JS divergence, while the blue solid line indicates the mean JS divergence. The upper and lower boundaries correspond to the upper and lower quartiles, respectively. The whiskers extend to 1.5 times the interquartile range (IQR).

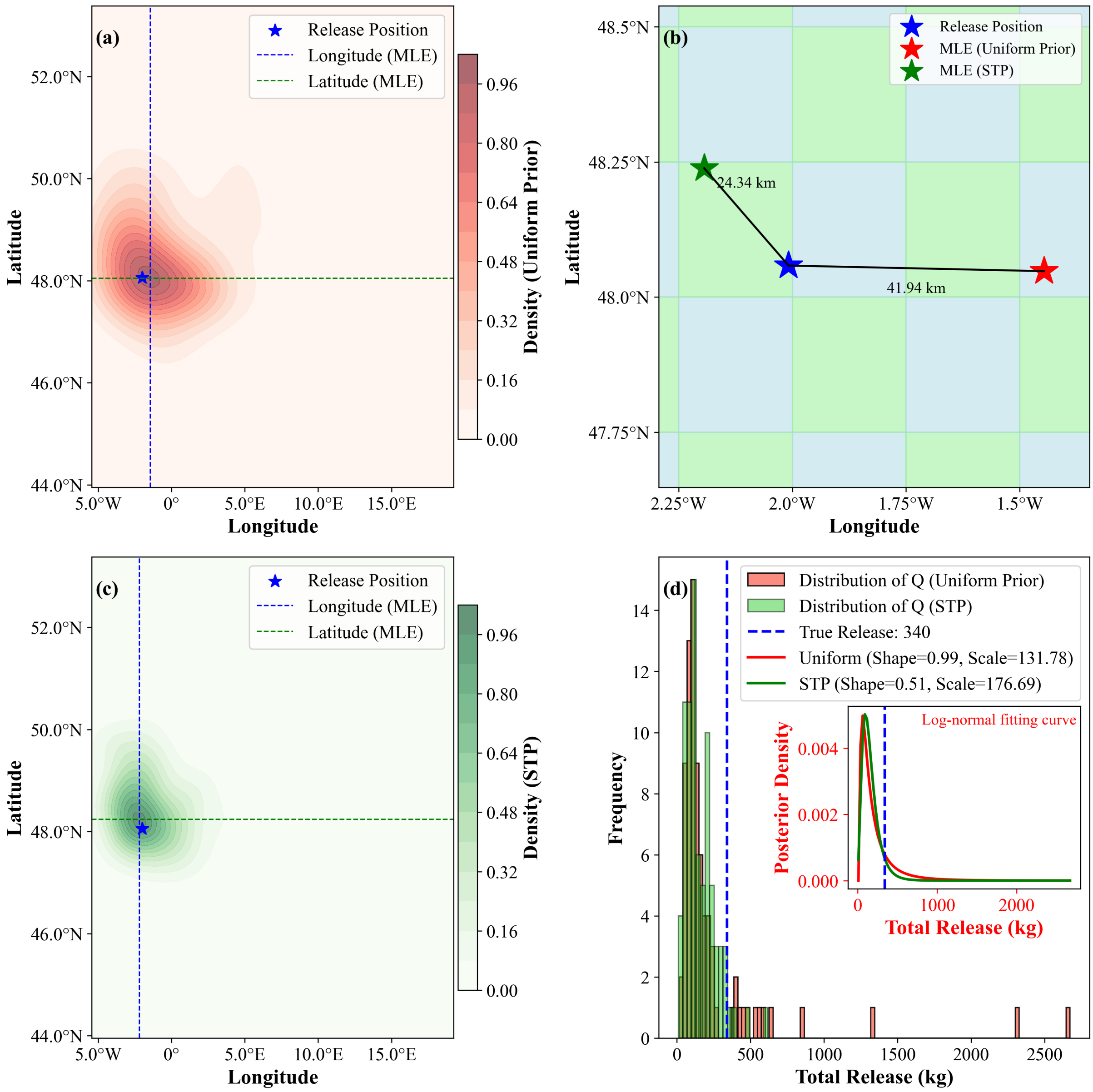


Fig. 7. Result's distribution of uniform prior and STP inversion after 80 runs. (a) Spatial distribution of inversion results based on the uniform prior. (b)Maximum likelihood estimation (MLE) of spatial distributions for uniform prior and STP inversion results, the grid represents the minimum resolution grid (0.25°) used in the dispersion simulation. (c) Spatial distribution of STP inversion results. (d) Inversion total release distributions for uniform prior and STP.

### 3. *Statistical evaluation of inversion results*

A comparison of Fig. 7a and 7c clearly reveals that the source positions obtained through STP inversion are more concentrated around the true source positions. Furthermore, Fig 7b illustrates the relative positions of the maximum likelihood estimates (MLE) of the inversion results for the uniform prior and STP, in comparison with the true release locations. It is evident that the relative distance between the MLE of STP and the true release location is significantly smaller than that for the uniform prior. Notably, the MLE obtained with STP lies within the same minimum resolution grid (0.25°) as the true release location. Fig. 7d illustrates the distribution of total release in the inversion results under two different prior conditions. It is evident that both priors underestimate the true release quantity, likely due to systematic biases in the input data. However, a closer examination of the distributions reveals a stark contrast: the inversion results with the uniform prior display numerous extreme values, with many inversion results significantly exceeding the true release quantity, despite the overall inversion being smaller. This suggests that using the traditional uniform prior is highly prone to instability in the inversion process. In contrast, the STP-based inversion results are more concentrated, with a notably

higher frequency near the true value. Since the prior distribution assumes that the release quantity follow a log-normal distribution, the results are fitted using a log-normal distribution. The STP inversion results are indeed more concentrated (with a smaller shape value) and closer to the true values (with a larger scale). Fig 8 illustrates the statistical distribution of relative errors in source inversion for the ETEX-I experiment, comparing the Bayesian method based on a STP with the traditional Bayesian method that employs a uniform prior. It is evident that the three types of relative errors in the STP inversion results are all smaller than those of the uniform prior. Specifically, the mean values (blue line) of the position relative error, total release relative error, and total relative error have been reduced by 25.9%, 27.7%, and 27.8%, respectively. The above results clearly demonstrate that STP effectively constrains the variance of the inversion results while enhancing their accuracy.

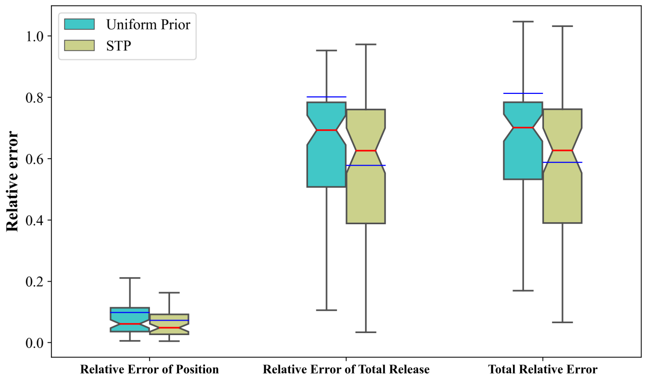


Fig. 8. Relative error distribution of source inversion results for ETEX-I experiment. The blue line represents the mean, while the red line denotes the median. The upper and lower boundaries correspond to the upper and lower quartiles, respectively. The whiskers extend to 1.5 times the interquartile range (IQR).

### 4. *Sensitivity of STP to the relative errors*

It is evident from Fig 9 that STP consistently outperforms the uniform prior in providing more concentrated and accurate location estimates under varying input data errors. Furthermore, as the input errors gradually increase, this "correction" effect does not show any signs of diminishing. Fig 10 presents a comparison of the relative errors in position, relative errors in total release, and total relative error under these conditions, with trends mirroring those in Figure 8. It is noteworthy that even in the absence of any input data errors, STP still demonstrates significantly better inversion results than the uniform prior. This indicates that the additional information introduced by STP compensates for the deficiencies in information utilization inherent in traditional methods, leading to a tangible improvement in inversion accuracy. In all cases, STP demonstrates superior inversion results. These results reveal STP's ability to limit the uncertainty in inversion outcomes caused by relative errors in the input data. This capability provides more reliable source information under conditions where errors are difficult to estimate in real inversion scenarios, and such reliability can play a pivotal role in emergency decision-making.

## IV. CONCLUSIONS

In this work, a Bayesian inversion method based on the backward plume spatiotemporal trajectory prior (STP) and joint adaptive MCMC sampling is proposed. The STP leverages fundamental fact that the backward plumes will overlap at the spatiotemporal coordinate of the release source, extracting prior information previously overlooked by traditional inversion methods from limited detection data, thereby enhancing both the accuracy and reliability of Bayesian inversion results. The joint adaptive MCMC sampling method introduces a Joint Adaptive Transition Proposal (JATP) to the conventional PT-based MCMC sampling. This approach addresses the challenge of designing transition proposals during sampling of high-dimensional complex objective functions, while also mitigating the non-Markovian nature of a single adaptive transition proposal. It strikes a better balance between global exploration and local characterization in the sampling process, ultimately improving sampling efficiency. Additionally, this paper investigates the optimal temperature schedule for the PT algorithm in the inversion problem, achieving faster convergence.

Subsequently, several inversion schemes based on the ETEX-I experiment were designed (fixed-sites validation experiment, statistical validation experiment, ideal twin experiment) to optimize and validate the proposed approach, while also investigating the sensitivity of the method to input data errors.

Firstly, the weights of the JATP were optimized. In the fixed-sites validation experiment, the acceptance rate curves for individual inversions using three adaptive transition proposal algorithms were compared to determine the optimal weight for the JATP. Verification showed that the weighted JATP consistently maintained an acceptance rate close to the theoretically derived optimal value (0.234) throughout the inversion iterations, thereby ensuring the efficiency of the sampling process.

Secondly, inversion calculations were performed using different temperature schedules in the fixed-sites validation experiment. The sensitivity relationships between the average burn-in period and the parameters of the PT algorithm's temperature schedule (number of chains, temperature intervals, and total temperature) were obtained, leading to the identification of the optimal temperature schedule for this method which reduced the average burn-in period by a factor of 2.89 compared to single-chain MCMC sampling.

Thirdly, using the complete ETEX-I experimental data, we qualitatively demonstrated the indicative effect of STP on the source term range, and quantitatively explored STP's sensitivity to data volume as well as its improvement effect on inversion results under the statistical validation experiment through the bootstrap method. The maximum probability

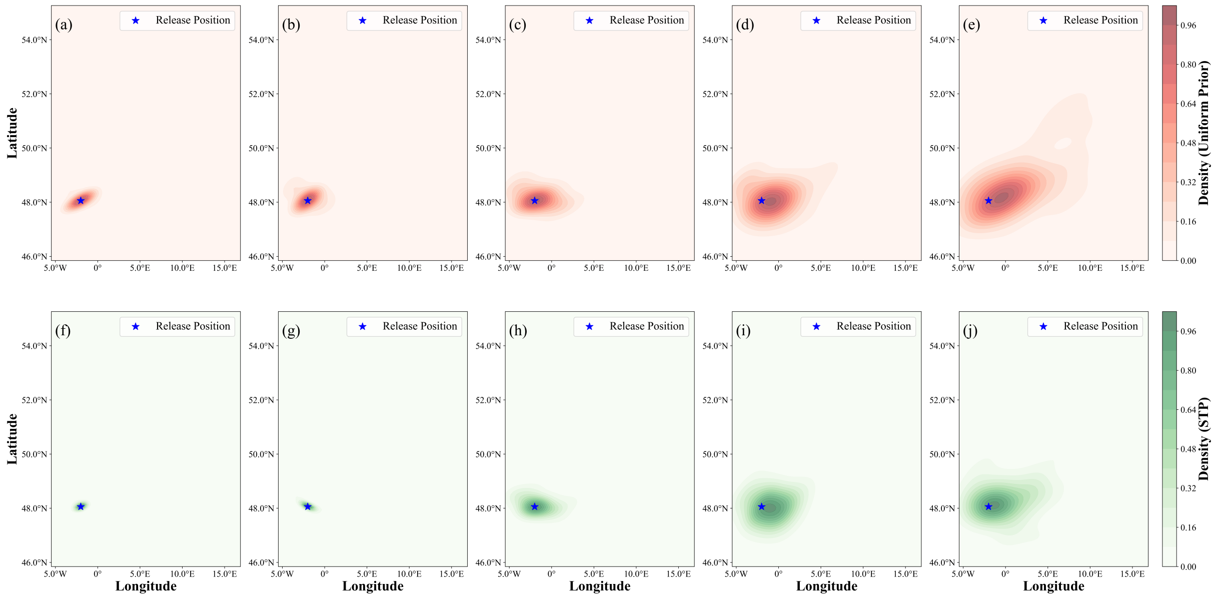


Fig. 9. Distribution of invertion source positions under varying levels of relative error. (a) and (b) represent 0% error, (c) and (d) represent 25% error, (e) and (f) represent 50% error, (g) and (h) represent 100% error, and (i) and (j) represent 200% error.

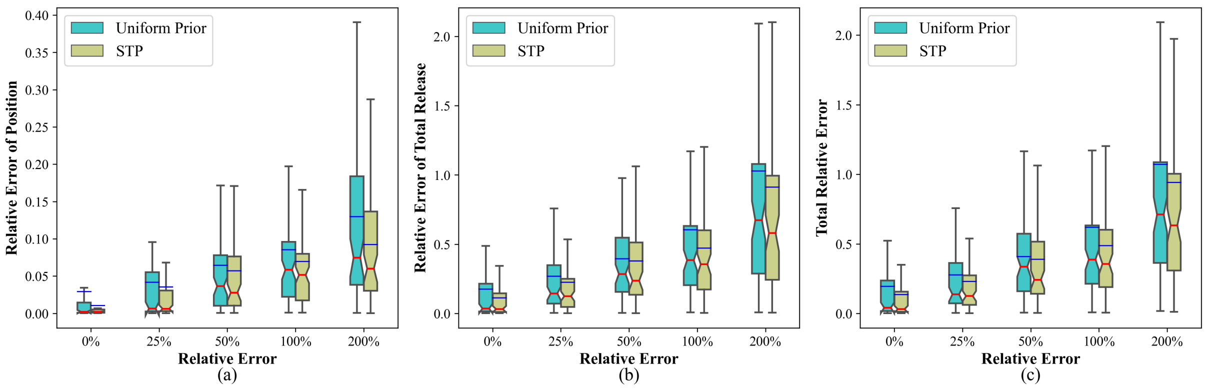


Fig. 10. The distribution of relative errors in the inversion results under different input data errors in the ideal twin experiment. (a) Relative error of position ($\delta_x$); (b) Relative error of total release ($\delta_Q$); (c) Total relative error ($\delta_{total}$).

density region of STP can relatively accurately estimate the location and timing of the release before the inversion calculation. Moreover, STP exhibits relatively low sensitivity to data volume, showing effective indication when the data set exceeds 30 points. Incorporating STP into the inversion process significantly enhances the accuracy of the inversion results while constraining the uncertainty range, thus yielding more reliable outcomes. In the ETEX-I case, the mean relative errors in position, release quantity, and total error were reduced by 25.9%, 27.7%, and 27.8%, respectively, compared to the traditional uniform prior. Meanwhile, STP reduced the maximum likelihood estimate of the invertion position distribution to within the smallest resolution (0.25°) grid.

Finally, using the ideal twin experiment, we investigated STP's sensitivity to input data errors by introducing different relative errors. The results demonstrated that STP consistently produces more accurate inversion results and smaller inversion error ranges under varying input data errors, with this improvement effect remaining strong even as the relative errors increase. It is also noteworthy that, even in the absence of any input data errors, STP still achieves higher inversion accuracy compared to the uniform prior.

These results indicate that, in practical scenarios, the optimized joint adaptive MCMC sampling and STP, when applied to Bayesian source inversion, can effectively enhance the inversion speed, accuracy, and robustness, thereby achieving satisfactory inversion results that meet the requirements of timeliness and accuracy in emergency decision-making. This method presents a promising solution for source inversion in atmospheric releases of radionuclides. However, the backward dispersion of the "plume" tends to cause the temporal prior to shift forward, resulting in an earlier inversion of the start time. Future work will focus on mitigating probability artifacts induced by dispersion to enhance the method further.

## Appendix A: Appendix A: Plots of the variable’s sampling trajectories with different temperature schedules

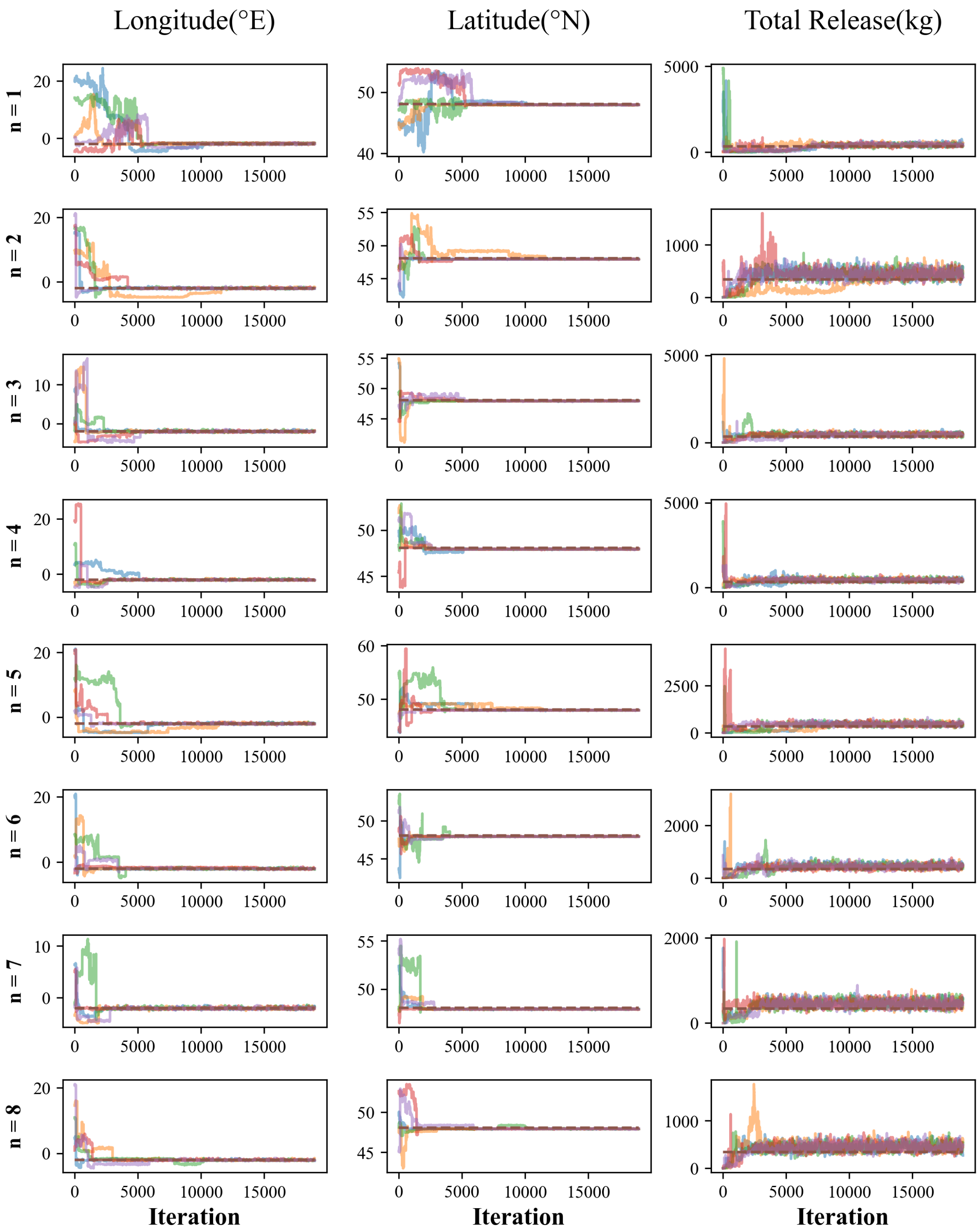


Fig. A.1. The sampling trajectories for the single-chain MCMC ($n = 1$) and the parallel tempering algorithm MCMC ($n = 2, 3, 4, 5, 6, 7, 8$) when the geometric spacing is $e$.The dashed line represents the true value of the variable and the solid lines represent the sampling curve.

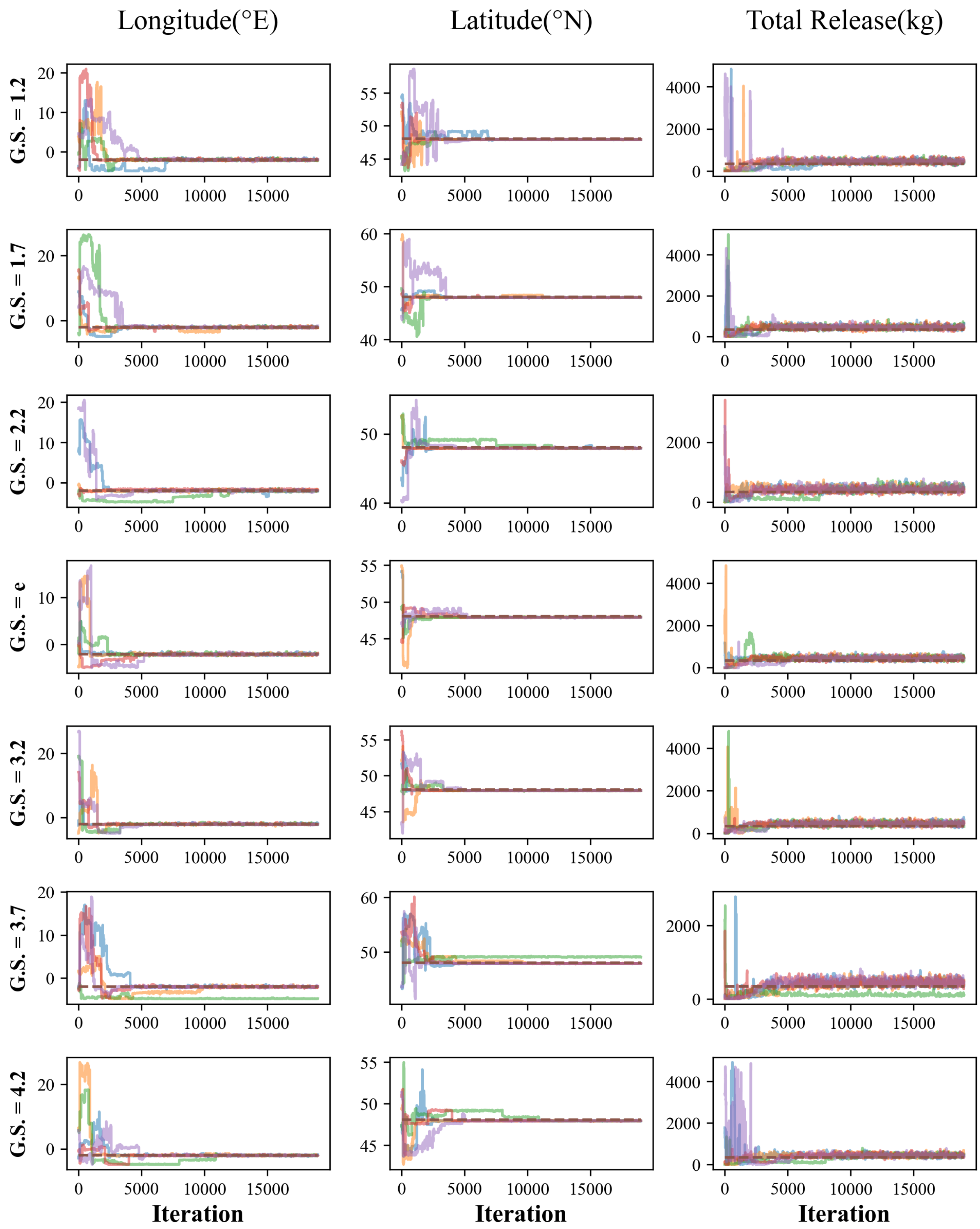


Fig. A.2. Plot of the variable's sampling trajectories at each geometric spacing (1.2, 1.7, 2.2, $e$, 3.2, 3.7, 4.2) with three chains. G. S. stands for geometric spacing.

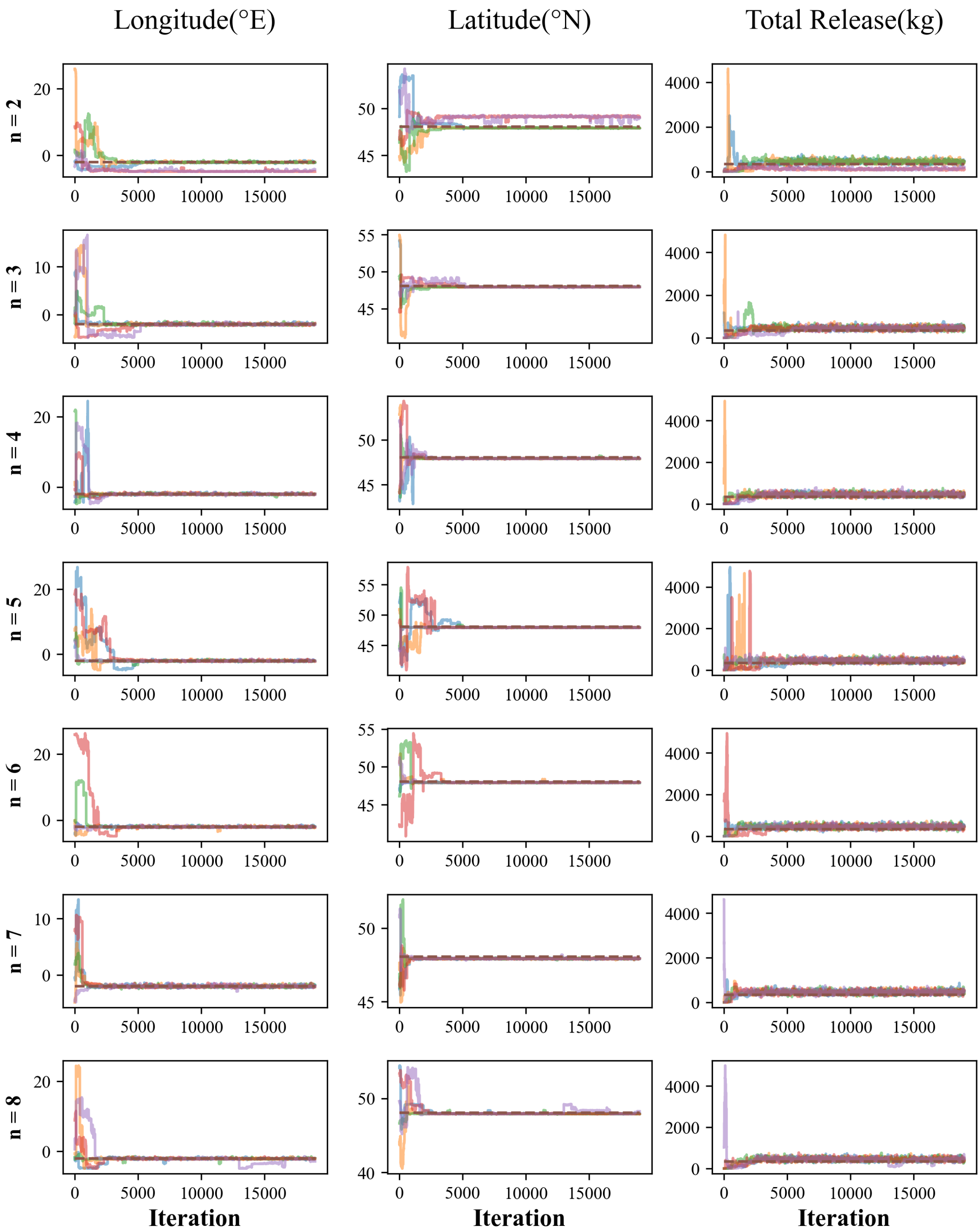


Fig. A.3. Plot of the variable's sampling trajectories at different chain numbers ($n = 2, 3, 4, 5, 6, 7, 8$) with a maximum temperature of $e^2$.

[1] L. R. Anspaugh, R. J. Catlin, M. Goldman, The global impact of the Chernobyl reactor accident. Science **242**, 1513-1519 (1988). doi:10.1126/science.3201240

[2] G. Katata, M. Ota, H. Terada et al., Atmospheric discharge and dispersion of radionuclides during the Fukushima Daiichi Nuclear Power Plant accident. Part I: Source term estimation and local-scale atmospheric dispersion in early phase of the accident. J. Environ. Radioactiv. **109**, 103-113 (2012). doi:10.1016/j.jenvrad.2012.02.006

[3] X. Zhang, W. Raskob, C. Landman et al., Sequential multi-nuclide emission rate estimation method based on gamma dose rate measurement for nuclear emergency management. J. Hazard. Mater. **325**, 288-300 (2017). doi:10.1016/j.jhazmat.2016.10.072

[4] O. Tichý, V. Šmídl, R. Hofman et al., Bayesian inverse modeling and source location of an unintended 131 I release in Europe in the fall of 2011. Atmos. Chem. Phys. **17**, 12677-12696 (2017). doi:10.5194/acp-17-12677-2017

[5] O. Masson, G. Steinhauser, D. Zok et al., Airborne concentrations and chemical considerations of radioactive ruthenium from an undeclared major nuclear release in 2017. P. Natl. Acad. SCI. USA. **116**, 16750-16759 (2019). doi:10.1073/pnas.1907571116

[6] J.-J. Ingremeau, O. Saunier, Investigations on the source term of the detection of radionuclides in North of Europe in June 2020. EPJ Nuc. Sci. Technol. **8**, 10 (2022). doi:10.1051/epjn/2022003

[7] Y. Xu, S. Fang, X. Dong el at., A spatiotemporally separated framework for reconstructing the sources of atmospheric radionuclide releases. Geosci. Model Dev. **17**, 4961–4982 (2024). doi:10.5194/gmd-17-4961-2024

[8] Y. Ling, Q. Yue, T. Huang et al., Multi-nuclide source term estimation method for severe nuclear accidents from sequential gamma dose rate based on a recurrent neural network. J. Hazard. Mater. **414**, 125546 (2021). doi:10.1016/j.jhazmat.2021.125546

[9] J. Wang, B. Wang, J. Liu et al., An inverse method to estimate the source term of atmospheric pollutant releases. Atmos. Environ. **260**, 118554 (2021). doi:10.1016/j.atmosenv.2021.118554

[10] J. Wang, R. Zhang, Z. Xin, Locating unknown number of multi-point hazardous gas leaks using Principal Component Analysis and a Modified Genetic Algorithm. Atmos. Environ. **230**, 117515 (2020). doi:10.1016/j.atmosenv.2020.117515

[11] X. L. Zhang, G. F. Su, J. G. Chen et al., Iterative ensemble Kalman filter for atmospheric dispersion in nuclear accidents: An application to Kincaid tracer experiment. J. Hazard. Mater. **297**, 329–339 (2015). doi:10.1016/j.jhazmat.2015.05.035

[12] X. Zheng, Z. Chen, Back-calculation of the strength and location of hazardous materials releases using the pattern search method. J. Hazard. Mater. **183**, 474-481 (2010). doi:10.1016/j.jhazmat.2010.07.048

[13] S. Fang, X. Dong, S. Zhuang et al., Oscillation-free source term inversion of atmospheric radionuclide releases with joint model bias corrections and non-smooth competing priors. J. Hazard. Mater. **440**, 129806 (2022). doi:10.1016/j.jhazmat.2022.129806

[14] O. Tichý, V. Šmídl, R. Hofman et al., LS-APC v1. 0: a tuning-free method for the linear inverse problem and its application to source-term determination. Geosci. Model Dev. **9**, 4297-4311 (2016). doi:10.5194/gmd-9-4297-2016

[15] J. M. Tomas, V. Peereboom, A. Kloosterman et al., Detection of radioactivity of unknown origin: Protective actions based on inverse modelling. J. Environ. Radioactiv. **235**, 106643 (2021). doi:10.1016/j.jenvrad.2021.106643

[16] D. Ma, W. Tan, Q. Wang et al., Location of contaminant emission source in atmosphere based on optimal correlated matching of concentration distribution. Process Saf. Environ. **117**, 498-510 (2021). doi:10.1016/j.psep.2018.05.028

[17] L. Pang, Y. Zhang, H. Qu et al., Approach to identifying pollutant source and matching flow field. Atmos. Environ. **73**, 1-10 (2013). doi:10.1016/j.atmosenv.2013.02.042

[18] O. Saunier, D. Didier, A. Mathieu et al., Atmospheric modeling and source inversion of radioactive ruthenium from an undeclared major release in 2017. P. Natl. Acad. SCI. USA. **116**, 24991-25000 (2019). doi:10.1073/pnas.1907823116

[19] J. H. Sørensen, Method for source localization proposed and applied to the October 2017 case of atmospheric dispersion of Ru-106. J. Environ. Radioactiv. **189**, 221-226 (2018). doi:10.1016/j.jenvrad.2018.03.010

[20] Y. Xue, X. Cui, K. Li et al., Statistical source analysis of recurring sulfur dioxide pollution events in a chemical industrial park. Atmos. Environ. **296**, 119564 (2023). doi:10.1016/j.atmosenv.2022.119564

[21] S. Guo, R. Yang, H. Zhang et al., Source identification for unsteady atmospheric dispersion of hazardous materials using Markov Chain Monte Carlo method. Int. J. Heat Mass Tran. **52**, 3955-3962 (2009). doi:10.1016/j.ijheatmasstransfer.2009.03.028

[22] Y. Wang, H. Huang, L. Huang et al., Evaluation of Bayesian source estimation methods with Prairie Grass observations and Gaussian plume model: A comparison of likelihood functions and distance measures. Atmos. Environ. **152**, 519–530 (2017). doi:10.1016/j.atmosenv.2017.01.014

[23] P. De Meutter, I. Hoffman, Bayesian source inversion of an anomalous Selenium-75 release at a nuclear research institute. J. Environ. Radioactiv. **218**, 106225-106225 (2020). doi:10.1016/j.jenvrad.2020.106225

[24] P. De Meutter, I. Hoffman, K. Ungar, On the model uncertainties in Bayesian source inversion using an ensemble of weather predictions, the emission inverse modelling system FREAR v1.0, and the Lagrangian transport and dispersion model Flexpart v9.0.2. Geosci. Model Dev. **14**, 1237-1252 (2021). doi:10.5194/gmd-14-1237-2021

[25] J. Dumont Le Brazidec, M. Bocquet, O. Saunier et al., MCMC methods applied to the inversion of the autumn 2017 Ruthenium-106 atmospheric contamination source. Atmos. Environ. **6**, 100071 (2020). doi:10.1016/j.aeaoa.2020.100071

[26] J. Dumont Le Brazidec, M. Bocquet, O. Saunier et al., Quantification of uncertainties in the assessment of an atmospheric release source applied to the autumn 2017 $^{106}$Ru event. Atmos. Chem. Phys. **21**, 13247 (2021). doi:10.5194/acp-21-13247-2021

[27] S. Qiu, B. Chen, R. Wang et al., Atmospheric dispersion prediction and source estimation of hazardous gas using artificial neural network, particle swarm optimization and expectation maximization. Atmos. Environ. **178**, 158-163 (2018). doi:10.1016/j.atmosenv.2018.01.056

[28] F. Xue, X. Li, W. Zhang, Bayesian identification of a single tracer source in an urban-like environment using a deterministic approach. Atmos. Environ. **164**, 128-138 (2017). doi:10.1016/j.atmosenv.2017.05.046

[29] Y. Zhao, Y. Liu, L. Wang et al., Source inversion of At-

mospheric Releases by Bayesian Inference and the Backward Atmospheric Dispersion Model: An Application to ETEX-I Data. Sci. Technol. Nucl. Ins. **2021**, 1-10 (2021). doi:10.1155/2021/5558825

[30] X. Zheng, Z. Chen, Inverse calculation approaches for source determination in hazardous chemical releases. J. Loss Prevent. Proc. **24**, 293-301 (2011). doi:10.1016/j.jlp.2011.01.002

[31] W. Cui, B. Cao, Q. Fan et al., Source term inversion of nuclear accident based on deep feedforward neural network. Ann. Nucl. Energy **175**, 109257 (2022). doi:10.1016/j.anucene.2022.109257

[32] D. Ma, Z. Zhang, Contaminant dispersion prediction and source estimation with integrated Gaussian-machine learning network model for point source emission in atmosphere. J. Hazard. Mater. **311**, 237–245 (2016). doi:10.1016/j.jhazmat.2016.03.022

[33] H. Van dop, R. Addis, G. Fraser et al., ETEX: A European tracer experiment; observations, dispersion modelling and emergency response. Atmos. Environ. **32**, 4089-4094 (1998). doi:10.1016/S1352-2310(98)00248-9

[34] T. Hesterberg, Bootstrap. Computation. Stat. **3**, 497-526 (2011). doi:10.1002/wics.182

[35] L. Adam, M. Branda, Sparse optimization for inverse problems in atmospheric modelling. Environ. Modell. Softw. **79**, 256-266 (2016). doi:10.1016/j.envsoft.2016.02.002

[36] S. Andronopoulos, I. V. Kovalets, Method of Source Identification Following an Accidental Release at an Unknown Location Using a Lagrangian Atmospheric Dispersion Model. ATMOSPHERE-BASEL **12**, 1305 (2021). doi:10.3390/atmos12101305

[37] H. Rajaona, F. Septier, P. Armand et al., An adaptive Bayesian inference algorithm to estimate the parameters of a hazardous atmospheric release. Atmos. Environ. **122**, 748-762 (2015). doi:10.1016/j.atmosenv.2015.10.026

[38] P. Seibert, A. Frank, Source-receptor matrix calculation with a Lagrangian particle dispersion model in backward mode. Atmos. Chem. Phys. **4**, 51-63 (2004). doi:10.5194/acp-4-51-2004

[39] I. Pisso, E. Sollum, H. Grythe et al., The Lagrangian particle dispersion model FLEXPART version 10.4. Geosci. Model Dev. **12**, 4955-4997 (2019). doi:10.5194/gmd-12-4955-2019

[40] A. Stohl, C. Forster, A. Frank et al., Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2. Atmos. Chem. Phys. **5**, 2461-2474 (2005). doi:10.5194/acp-5-2461-2005

[41] E. Al Aawar, S. El Mohtar, I. Lakkis et al., Bayesian source identification of urban-scale air pollution from point and field concentration measurements. Computat. Geosci. **27**, 605-626 (2023). doi:10.1007/s10596-023-10206-5

[42] F. Septier, P. Armand, C. Duchenne, Sequential Monte Carlo sampler applied to source term estimation in complex atmospheric environments. Atmos. Environ. **269**, 118822 (2022). doi:10.1016/j.atmosenv.2021.118822

[43] J. Zhu, X. Zhou, B. Cong et al., Estimation of the point source parameters by the adjoint equation in the time-varying atmospheric environment with unknown turn-on time. Build. Environ. **230**, 110029 (2023). doi:10.1016/j.buildenv.2023.110029

[44] Y. Liu, J. M. Haussaire, M. Bocquet et al., Uncertainty quantification of pollutant source retrieval: comparison of Bayesian methods with application to the Chernobyl and Fukushima Daiichi accidental releases of radionuclides. Q. J. Roy. Meteor. Soc. **143**, 2886-2901 (2017). doi:10.1002/qj.3138

[45] H. Haario, E. Saksman, J. Tamminen, An Adaptive Metropolis Algorithm. Bernoulli **7**, 223-242 (2001). doi:10.2307/3318737

[46] C. J. F. T. Braak, A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: easy Bayesian computing for real parameter spaces. Stat. Comput. **16**, 239-249 (2006). doi:10.1007/s11222-006-8769-1

[47] H. Haario, E. Saksman, J. Tamminen, Componentwise adaptation for high dimensional MCMC. Computation. Stat. **20**, 265-273 (2005). doi:10.1007/BF02789703

[48] G. O. Roberts, J. S. Rosenthal, Optimal Scaling for Various Metropolis-Hastings Algorithms. Stat. Sci. **16**, 351-367 (2001). doi:10.1214/ss/1015346320

[49] C. P. Robert, V. Elvira, N. Tawn et al., Accelerating MCMC algorithms. WIREs Comput. Stat. **10**, e1435-n/a (2018). doi:10.1002/wics.1435

[50] X. Zhang, C. Delpha, D. Diallo, Incipient fault detection and estimation based on Jensen–Shannon divergence in a data-driven approach. Signal Process. **169**, 107410 (2020). doi:10.1016/j.sigpro.2019.107410